

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/793,408DATE: 09/25/97  
TIME: 12:34:44

INPUT SET: S20569.raw

This Raw Listing contains the General  
Information Section and up to the first 5 pages.

## SEQUENCE LISTING

ENTERED

RECEIVED

NOV 04 1997

GROUP 1800

## (1) General Information:

(i) APPLICANT: Choo, Yen  
Klug, Aaron  
Sanchez Garcia, Isidro

(ii) TITLE OF INVENTION: Improvements in or Relating to  
Binding Proteins for Recognition of DNA

(iii) NUMBER OF SEQUENCES: 18

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
(B) STREET: 1100 New York Avenue, N.W.  
(C) CITY: Washington  
(D) STATE: D.C.  
(E) COUNTRY: USA  
(F) ZIP: 20005-3918

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: Word Perfect

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/GB95/01949  
(B) FILING DATE: 17-AUG-1995

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: GB 9514698.1  
(B) FILING DATE: 18-JUL-1995

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: GB 9422534.9  
(B) FILING DATE: 08-NOV-1994

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: GB 9416880.4

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/793,408DATE: 09/25/97  
TIME: 12:34:46

INPUT SET: S20569.raw

47 (B) FILING DATE: 20-AUG-1994

48

49 (2) INFORMATION FOR SEQ ID NO: 1:

50

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 60 base pairs

53 (B) TYPE: nucleic acid

54 (C) STRANDEDNESS: single

55 (D) TOPOLOGY: linear

56

57 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

58

59 CTCCTGCACT TGGACCTGTG CCATGGCCGG CTGGGCCGCA TAGAATGGAA

50

60 CAACTAAAGC

60

61

62 (2) INFORMATION FOR SEQ ID NO: 2:

63

64 (i) SEQUENCE CHARACTERISTICS:

65 (A) LENGTH: 92 amino acids

66 (B) TYPE: amino acid

67 (C) STRANDEDNESS:

68 (D) TOPOLOGY: unknown

69

70 (ii) MOLECULE TYPE: protein

71

72

73 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

74

75 Met Ala Glu Glu Arg Pro Tyr Ala Cys Pro

76

5

10

77

78 Val Glu Ser Cys Asp Arg Arg Phe Ser Arg

79

15

20

80

81 Ser Asp Glu Leu Thr Arg His Ile Arg Ile

82

25

30

83

84 His Thr Gly Gln Lys Pro Phe Gln Cys Arg

85

35

40

86

87 Ile Cys Met Arg Asn Phe Ser Xaa Xaa Xaa

88

45

50

89

90 Xaa Leu Xaa Xaa His Xaa Arg Thr His Thr

91

55

60

92

93 Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys

94

65

70

95

96 Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg

97

75

80

98

99 Lys Arg His Thr Lys Ile His Leu Arg Gln

PAGE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/793,408DATE: 09/25/97  
TIME: 12:34:48

INPUT SET: S20569.raw

100 85 90  
101  
102 Lys Asp  
103  
104 (2) INFORMATION FOR SEQ ID NO: 3:  
105  
106 (i) SEQUENCE CHARACTERISTICS:  
107 (A) LENGTH: 26 base pairs  
108 (B) TYPE: nucleic acid  
109 (C) STRANDEDNESS: single  
110 (D) TOPOLOGY: linear  
111  
112 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
113  
114 TATGACTTGG ATGGGAGACC GCCTGG 26  
115  
116 (2) INFORMATION FOR SEQ ID NO: 4:  
117  
118 (i) SEQUENCE CHARACTERISTICS:  
119 (A) LENGTH: 28 base pairs  
120 (B) TYPE: nucleic acid  
121 (C) STRANDEDNESS: single  
122 (D) TOPOLOGY: linear  
123  
124 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
125  
126 AATTCCAGGC GGTCTCCCAT CCAAGTCA 28  
127  
128 (2) INFORMATION FOR SEQ ID NO: 5:  
129  
130 (i) SEQUENCE CHARACTERISTICS:  
131 (A) LENGTH: 21 base pairs  
132 (B) TYPE: nucleic acid  
133 (C) STRANDEDNESS: single  
134 (D) TOPOLOGY: linear  
135  
136 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
137  
138 TATATAGCGT GGGCGTATAT A 21  
139  
140 (2) INFORMATION FOR SEQ ID NO: 6:  
141  
142 (i) SEQUENCE CHARACTERISTICS:  
143 (A) LENGTH: 24 base pairs  
144 (B) TYPE: nucleic acid  
145 (C) STRANDEDNESS: single  
146 (D) TOPOLOGY: linear  
147  
148 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
149  
150 GCGTATATAC GCCCAGCTA TATA 24  
151  
152 (2) INFORMATION FOR SEQ ID NO: 7:

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/793,408DATE: 09/25/97  
TIME: 12:34:51

INPUT SET: S20569.raw

153  
154 (i) SEQUENCE CHARACTERISTICS:  
155 (A) LENGTH: 21 base pairs  
156 (B) TYPE: nucleic acid  
157 (C) STRANDEDNESS: single  
158 (D) TOPOLOGY: linear  
159  
160 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
161  
162 TATATAGCGN NNGCGTATAT A 21  
163  
164 (2) INFORMATION FOR SEQ ID NO: 8:  
165  
166 (i) SEQUENCE CHARACTERISTICS:  
167 (A) LENGTH: 24 base pairs  
168 (B) TYPE: nucleic acid  
169 (C) STRANDEDNESS: single  
170 (D) TOPOLOGY: linear  
171  
172 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
173  
174 GCGTATATAC GCNNCGCTA TATA 24  
175  
176 (2) INFORMATION FOR SEQ ID NO: 9:  
177  
178 (i) SEQUENCE CHARACTERISTICS:  
179 (A) LENGTH: 33 base pairs  
180 (B) TYPE: nucleic acid  
181 (C) STRANDEDNESS: single  
182 (D) TOPOLOGY: linear  
183  
184  
185 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
186  
187 TTCCATGGAG ACGCAGAAGC CCTTCAGCGG CCA 33  
188  
189 (2) INFORMATION FOR SEQ ID NO: 10:  
190  
191 (i) SEQUENCE CHARACTERISTICS:  
192 (A) LENGTH: 33 base pairs  
193 (B) TYPE: nucleic acid  
194 (C) STRANDEDNESS: single  
195 (D) TOPOLOGY: linear  
196  
197 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
198  
199 TTCCATGGAG ACGCAGGTGA GTTCCTGACG CCA 33  
200  
201 (2) INFORMATION FOR SEQ ID NO: 11:  
202  
203 (i) SEQUENCE CHARACTERISTICS:  
204 (A) LENGTH: 33 base pairs  
205 (B) TYPE: nucleic acid

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/793,408DATE: 09/25/97  
TIME: 12:34:53

INPUT SET: S20569.raw

206 (C) STRANDEDNESS: single  
207 (D) TOPOLOGY: linear  
208

209 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
210

211 CCCCTTTCTC TTCCAGAAGC CCTTCAGCGG CCA  
212

33

213 (2) INFORMATION FOR SEQ ID NO: 12:  
214

215 (i) SEQUENCE CHARACTERISTICS:

216 (A) LENGTH: 33 amino acids

217 (B) TYPE: amino acid

218 (C) STRANDEDNESS:

219 (D) TOPOLOGY: unknown  
220

221 (ii) MOLECULE TYPE: peptide  
222

223 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
224

225 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg  
226 5 10  
227

228 Ile Cys Met Arg Asn Phe Ser Asp Arg Ser  
229 15 20  
230

231 Ser Leu Thr Arg His Thr Arg His Thr Gly  
232 25 30  
233

234 Glu Lys Pro  
235

236 (2) INFORMATION FOR SEQ ID NO: 13:  
237

238 (i) SEQUENCE CHARACTERISTICS:

239 (A) LENGTH: 33 amino acids

240 (B) TYPE: amino acid

241 (C) STRANDEDNESS:

242 (D) TOPOLOGY: unknown  
243

244 (ii) MOLECULE TYPE: peptide  
245

246 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
247

248 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg  
249 5 10  
250

251 Ile Cys Met Arg Asn Phe Ser Glu Arg Gly  
252 15 20  
253

254 Thr Leu Ala Arg His Glu Lys His Thr Gly  
255 25 30  
256

257 Glu Lys Pro  
258

PAGE: 1

**SEQUENCE VERIFICATION REPORT**  
**PATENT APPLICATION US/08/793,408**

DATE: 09/25/97  
TIME: 12:34:56

*INPUT SET: S20569.raw*

Line

Error

Original Text